

Continuous belief functions to qualify sensors performances

Pierre-Emmanuel Doré¹, Christophe Osswald¹, and Arnaud Martin²

¹ E^3I^2 , ENSTA Bretagne, 2 rue François Verny 29806 Brest Cedex 9, France

² IRISA, Université de Rennes 1, Lannion, France

{pierre-emmanuel.dore,christophe.osswald}@ensta-bretagne.fr
arnaud.martin@univ-rennes1.fr

Abstract. In this paper, we deal with the problem of sensor performance estimation. As we assume that the sensor is described with only few data, we decide to use the theory of belief functions to represent the inherent uncertainty of our information. Hence, we introduce the belief functions framework, especially in the continuous approach. We describe the model of sensor adopted in our study. Knowing the experimental setting, we suggest an approach to model the sources of information describing our sensor. Finally, we combine these sources in order to estimate sensor performances.

Keywords: Sensor performances, Continuous belief function, Parametric model, Inference, Fusion

1 Introduction

The theory of belief functions has been introduced by the famous works of A. Dempster about upper and lower probabilities [3,4] and those of G. Shafer [13] on the theory of evidence. The work of Ph. Smets [14] contributed to spread it in the scientific community. Recently, thanks to new breakthroughs [18,12,15,16,2], the application of belief functions on continuous framework has gained some interest. Hence, we apply these results to describe sensor performances.

Wireless sensors networks are more and more used in monitoring applications [11,9]. One crucial issue in this domain is the sensors placement. Indeed, the aim is to place the sensors in order to maximize the chance to detect an intrusion. To fulfil this objective, we have to characterize the performances of a sensor as a detector. To estimate the performance of a sensor, we adopt a parametric approach. As we only have a small amount of measures to define it, we decide to use the belief function framework to take into account the lack of information.

In a first part (section 2), we present the theory of belief function within a continuous frame of discernement. Then, we introduce a parametric model describing sensor performances (section 3). As we have few data, we present some results obtained by A. Dempster and we suggest an approach funded on the maximum necessity and likelihood principles to the estimate the parameters describing the sensor performances (section 4). Hence, considering that the

experimental settings provide us three cognitive independant sources of informations (we take measures in three different places), we merge their respective beliefs about the value of the sensor parameters (section 5). Finally, we analyse our results in section 6. Thanks to these operations, we are able to characterize more accurately the sensor performances.

2 Belief function framework

The theory of belief functions is a tool used to represent the imperfection of a source of information. There are many kinds of imperfection a belief function can describe such as ignorant, vagueness, uncertainty, ... The purpose of this section is to recall some useful parts of this theory.

2.1 Belief on real numbers

In [15], Ph. Smets describes an approach of belief functions on real number. He suggests to assign mass on the intervals of $\mathbb{R} = \mathbb{R} \cup \{-\infty, \infty\}$. There is a lot of advantage to procede in this way. Indeed, we can easily associate a basic belief density on \mathbb{R} to a probability density function on \mathbb{R}^2 . However, this framework is quite restrictive and we cannot use it to describe belief function with basic belief assignement on unconnected sets. In a previous work [6], we suggest to scan the set of focal elements (the subset of \mathbb{R} whose the basic belief assignement is not null), \mathcal{F} , using an index function f and a specific index space I .

$$\begin{aligned} f^I : I &\longrightarrow \mathcal{F} \\ y &\longmapsto f^I(y) \end{aligned} \tag{1}$$

With this index function, we consider a positive measure μ^{Ω} ¹ such as $\int_I d\mu^{\Omega}(y) \leq 1$.

Hence, to define a belief function, we have to consider the brace (f^I, μ^{Ω}) . In order to compute belief functions, we need to define for all A in $\mathcal{P}(\Omega)$ (a family of subset of Ω):

$$F_{\subseteq A} = \{y \in I \mid f^I(y) \subseteq A\} \tag{2}$$

$$F_{\cap A} = \{y \in I \mid (f^I(y) \cap A) \neq \emptyset\} \tag{3}$$

$$F_{\supseteq A} = \{y \in I \mid A \subseteq f^I(y)\} \tag{4}$$

Once these subsets are defined, we can compute the following belief functions:

- The belief function:

$$bel^{\Omega}(A) = \int_{F_{\subseteq A}} d\mu^{\Omega}(y) \tag{5}$$

¹ For the sake of simplicity, it happens we considers that $d\mu^{\Omega}(y) = m^{\Omega}(f^I(y))$.

– The plausibility function:

$$pl^\Omega(A) = \int_{F_{\cap A}} d\mu^\Omega(y) \quad (6)$$

– The communality function:

$$q^\Omega(A) = \int_{F_{\supseteq A}} d\mu^\Omega(y) \quad (7)$$

In this framework, we define some basic tools. One of them is the conjunctive rule of combination which allow us to merge sources of information. Let (f_1^I, μ_1^Ω) and (f_2^I, μ_2^Ω) two belief functions. The conjunctive rule of combination [17] brings the belief function $(f_1^I \odot_2, \mu_1^\Omega \odot_2)$ [6] such that:

$$q_1^\Omega \odot_2(A) = q_1^\Omega(A) \cdot q_2^\Omega(A) \quad (8)$$

Within this framework, we will study a particular type of belief functions, the consonant ones.

A consonant belief function is a belief function whose the focal sets are nested. This allows us to create a total ordering on \mathcal{F} linked to the \subseteq relation. Hence, we can define an index function f from \mathbb{R}^+ , to \mathcal{F} such as $(y \geq x) \implies (f(y) \subseteq f(x))$ [15]. To generate consonant sets, we can use g , a continuous function from \mathbb{R}^n to $I = [0, \alpha_{max}] \subset \mathbb{R}^+$. The α -cuts of g are the set such as:

$$f_{cs}^I(\alpha) = \{x \in \mathbb{R}^n | g(x) \geq \alpha\} \quad (9)$$

We have the property that $F_{\subseteq A}^{cs}$ is an element of Borel algebra. Indeed:

$$\begin{aligned} F_{\subseteq A}^{cs} \neq \emptyset &\implies \exists \alpha_{\inf} = \inf \{\alpha \in I | f_{cs}^I(\alpha) \cap A \neq \emptyset\} \\ &\implies F_{\subseteq A}^{cs} = [\alpha_{\inf}, \alpha_{\max}] \end{aligned} \quad (10)$$

Using a similar argument, we can prove that $F_{\supseteq A}^{cs}$ and $F_{\cap A}^{cs}$ are elements of Borel algebra. Hence, we can define the index function:

$$\begin{aligned} f_{cs}^I : I = [0, \alpha_{\max}] &\longrightarrow \{f_{cs}^I(\alpha) | \alpha \in I\} \\ \alpha &\longmapsto f_{cs}^I(\alpha) \end{aligned} \quad (11)$$

If we consider a probability measure $\mu^{\mathbb{R}^n}$ on \mathcal{I} , the brace $(f_{cs}^I, \mu^{\mathbb{R}^n})$ refers to a belief function.

2.2 Maximum of necessity

When we work with an “objective” source of information, we can apply the principle of maximum of necessity. This principle comes from the theory of possibility [7,8,10]. The theory of possibility is a particular case of the theory of belief functions (focal elements are always nested). In this situation, the plausibility corresponds to possibility distribution and we consider the necessity which

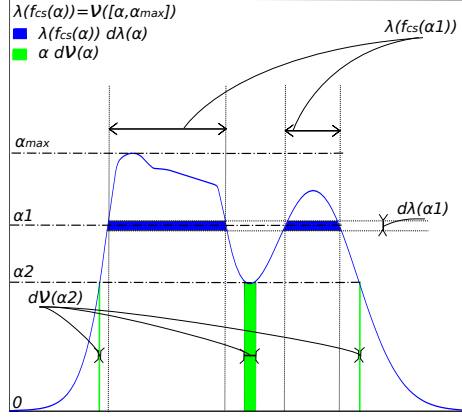


Fig. 1. How to build belief functions from a probability density function $Betf$.

corresponds to the belief function. The idea is to work with the most informative distribution of possibility (for the necessity ordering) which fulfils the following assumptions. The first one is that the possibility dominates the probability, *i.e.* for all A measurable $\Pi(A) \geq P(A)$. The second one is that the ordering must be kept, *i.e.* $P(A) \geq P(A') \Leftrightarrow \Pi(A) \geq \Pi(A')$.

These conditions can be transposed in the framework of belief functions [10]. Finding a belief function which verifies these properties is equivalent to find a nested focal sets family such as for all A belonging to this family, A is the smallest set (for the inclusion ordering) such as $P(A) = \beta$. This sets family corresponds to the confidence sets in theory of probability. If we have as input a continuous probability density function $Betf$, the focal set $(f_{cs}^I(\alpha))$ is the α -cuts of $Betf$. We obtain a belief function defined by $(f_{cs}^I, \mu^B(\mathbb{R}))$ such as if we adapt the result of [10], we obtain [6] (*cf.* figure 1):

$$pl^B(\mathbb{R})(x) = 1 - BetP(f_{cs}^I(Betf(x))) \quad (12)$$

*i.e.*²:

$$d\mu^R(\alpha) = \alpha dV(\alpha) \text{ with } V([\alpha, \alpha_{\max}]) = \lambda(f_{cs}^I(\alpha)) \quad (13)$$

Within this framework, we can build belief functions on real number with complex focal sets. In most of the cases, the information given by a sensor is represented with a probability distribution. However, it happens we do not have enough information to precisely define a distribution a probability. A belief function can be useful to take this phenomenon into account.

² λ refers to the Lebegue's measure.

3 Model of sensor

We consider in this paper that a sensor is a detector. Each detector is defined using two characteristic measures, the probability of true positive (P_t) and the probability of false positive (P_f). The probability of true positive is the probability that we detect an object when there is actually something. The probability of false positive is the probability we decide there is something when there is nothing. In the literature, the probability of true positive has been defined as a function of the distance between the sensor and the object we want to detect. In our study case, to model the sensor, we make the following assumptions:

1. The sensor is passive. Hence, the probability of false positive is not a function of the distance between the sensor and the object. When the distance between the sensor and the object is growing, the value of P_t tends to P_f .
2. If the distance between the sensor and the object is smaller than α , the detection is guaranteed.
3. Beyond this point, the probability of true positive decreases such as the inverse of a geometric law of parameter λ . This type of behaviour has been observed on seismic and magnetic sensor or microphone.

Hence, the following equation describes the probability of detection according a distance d between the sensor and the target, the parameters α , λ and P_f :

$$P_t(d) = \begin{cases} 1 & \text{if } d \leq \alpha \\ \alpha^\lambda \cdot \frac{1 - P_f}{d^\lambda} + P_f & \text{otherwise} \end{cases} \quad (14)$$

This is a trade off between the Elfe's model [9] and the geometric model of sensors [11]. The shape of the sensor coverage (the probability of true positive according the distance) described in equation (14) is represented in figure 2.

We will estimate the coverage of a sensor, using our parametric model, when we dispose of only a small amount of measures.

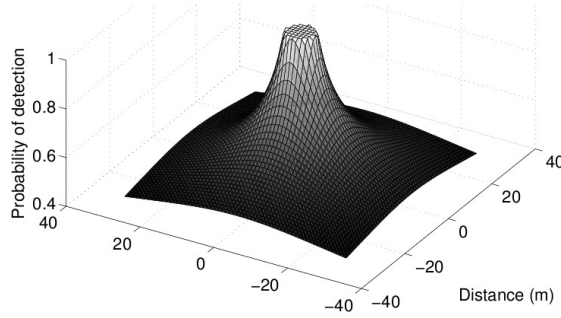


Fig. 2. Probability of detection on a grid. The sensor position defines the origin of our coordinates ($\alpha = 4$, $\lambda = 0.5$, $P_f = 0.15$).

4 Belief functions induced by sampling

The size of the sample set is not large enough to define precisely a distribution of probability. As measure prospecting is an expensive process, we have to find a way to describe imprecise knowledge on true and false positive probabilities using a small sample set. We will apply our work to the estimation of the sensor performances.

4.1 Experimental settings

In order to illustrate our work, we assume we monitor vehicles moving on three roads at different distances of the sensor (*cf.* figure 3). In each case, we have

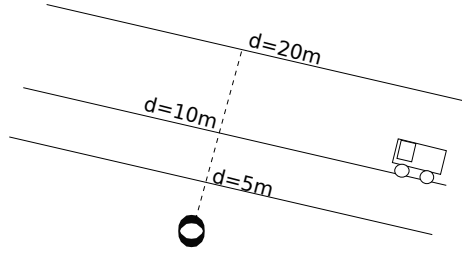


Fig. 3. The experimental setting.

obtained a different empirical value of the probability of detection according the distance. The generic values used in our work are described in table 1.

Distance sensor/vehicule d	number of test n	number of detection k
5 m	20	18
10 m	20	14
20 m	20	10

Table 1. Generic values.

As we work on a small amount of data, the estimates of probabilities of true positive are imprecise. In fact, a family of probability distributions can fit with the data we obtain. The theory of belief functions is an efficient framework to represent a family of probability distributions. Hence, we present some methods to model information induced by a small learning set in the belief functions framework.

4.2 The Dempster's approach

The first motive of Dempster [5] when he developped the theory of upper and lower probabilities was to propose a mathematical framework to model statistical inference. One of his main results in this domain concerns the binomial sampling. When we have a small amount of samples following a binomial law, it is impossible to precizely define the parameter p of this law. Hence we define a set of distributions of probability using belief functions theory. Let \mathcal{P} be the interval $[0, 1]$ which describes the parameter p of a binomial law. Let n be the size of the sample set and k be the number of samples belonging to one of the two classes of the binomial sampling. We obtain a conditional belief function $m^{\mathcal{P}}$ on the value of p over \mathcal{P} knowing n and k [5] :

$$\begin{aligned} m^{\mathcal{P}}[n, k]([u, v]) &= \frac{n!}{(k-1)!(n-k-1)!} u^{k-1} (1-v)^{n-k-1} & 0 < k < n \\ m^{\mathcal{P}}[n, k]([0, v]) &= n(1-v)^{n-1} & k = 0 \\ m^{\mathcal{P}}[n, k]([u, 1]) &= nu^{n-1} & k = n \end{aligned} \quad (15)$$

This kind of belief function can model our problem and be used to define our probabilities. Unfortunately, they are difficult to handle and we cannot use on them the *max* or the *min* operators. Indeed, generally, these operators do not induce belief functions (*cf.* example in table 2). Hence, we have to find another way to characterize our information. There are methods using confidence regions or least commitment principle [1]. However, in this paper we focus on the maximum of necessity principle as it will sum up the information contained by all the confidence regions in an unique function.

focal elements	\emptyset	a	b	c	ab	ac	bc	abc
$m1$	0	0.25	0.11	0.03	0.06	0.10	0.15	0.3
$m2$	0	0.22	0.07	0.04	0.09	0.10	0.13	0.35
$pl1$	0	0.68	0.59	0.94	0.61	0.89	0.75	1
$pl2$	0	0.71	0.59	0.91	0.67	0.93	0.78	1
$pl_{\max} = \max(pl1, pl2)$	0	0.71	0.59	0.94	0.67	0.93	0.78	1
m_{\max}	0	0.415	0.1723	0.0040	0.1446	0.0923	0.2003	-0.0285

Table 2. The max operator does not induce a belief function.

4.3 Likelihood and maximum of necessity principles

As we cannot apply *max* operator on the belief function defined in equation (15), we decide to use the maximum likelihood principle [14] to define a belief function associated to a binomial sampling. This principle can be resumed as following:

Assuming a likelihood function l , the likelihood of a parameter θ knowing that x is true is equal to the likelihood of x knowing that θ is true.

$$l[x](\theta) = l[\theta](x) \quad (16)$$

The likelihood function l must satisfy several properties. One of them is that the likelihood of a set of hypothesis is higher than the likelihood of its subsets:

$$l(\Theta) \geq l(\theta), \theta \subseteq \Theta \quad (17)$$

Another one is that the likelihood function must be a sub-additive function:

$$l(\theta_A \cup \theta_B) \leq l(\theta_A) + l(\theta_B) \quad (18)$$

Indeed, two disjoint sets of hypothesis can have a likelihood equals to 1. Hence, we can use the plausibility as a likelihood function. In our study, according the likelihood principle, we obtain for π subset of \mathcal{P} and k in $\llbracket 0, n \rrbracket$:

$$pl[n, k](\pi) = pl[n, \pi](k) \quad (19)$$

Hence we have to define $pl[n, \pi](k)$. The distribution of probability P describing a binomial law of parameter p is equal to :

$$P[n, p](k) = \binom{n}{k} p^k (1-p)^{n-k}, 0 \leq k \leq n \quad (20)$$

Applying the maximum of necessity principle to this distribution of probability [8], we define a consonant continuous belief function $pl[n, p]$. This one is entirely described by his contour function computed using algorithm 1. In parametric models, data are highly dependant, hence we prefer not to use a conjunctive rule to merge source of information but a disjunctive one:

$$pl[n, k](\pi) = pl[n, \pi](k) = \max_{p \in \pi} (pl[n, p](k)) \quad (21)$$

Result: Plausibility of p for a given pair (n, k) .

$tmp = 0$;

$$M = \binom{n}{k} p^k (1-p)^{n-k};$$

for $i \in [0, n]$ **do**

$$P[n, p](i) = \binom{n}{i} p^i (1-p)^{n-i};$$

if $P[n, p](i) \leq M$ **then**

$$| \quad tmp = tmp + P[n, p](i);$$

end

end

$pl[n, p](k) = tmp$; // maximum of necessity principle.

$pl[n, k](p) = pl[n, p](k)$; // likelihood principle.

Algorithm 1: How to compute plausibility of p knowing (n, k) .

5 Combination of information

We have previously defined belief functions to describe the plausibility of P_t and P_f for given values of d . As we want to define these probabilities on the whole space, we have, using the model of sensor described by equation (14), to infer the plausibility of the 3-tuplet (α, k, P_f) .

5.1 Independant sources of information

Using the sensor model (*cf.* equation (14)), we can associate for each distance d of measure a 3-tuplet (α, k, P_f) to a given probability of true positive P_t . Moreover, using table 1, we can define $pl_d^{\mathcal{P}}(P_t)$. Let \mathcal{T} be the parameter framework. We obtain thanks to a mass transfert:

$$pl_d^{\mathcal{T}}(\alpha, \lambda, P_f) = pl_d^{\mathcal{P}}(P_t) \quad (22)$$

The experiments which lead us to define P_t for given distances d are cognitively independent. Hence we can merge these different sources of information using the conjunctive rule of combination[17]. As we use consonant belief functions, this operation can be written:

$$pl^{\mathcal{T}}(\alpha, \lambda, P_f) = \prod_{d \in \{5, 10, 20\}} pl_d^{\mathcal{T}}(\alpha, \lambda, P_f) \quad (23)$$

Hence, we assume that all the sources of information are relevant and reliable.

5.2 Inference using the max

For a given distance d and probability of true positive (P_t), there is, according the sensor model described in equation (14), a set \mathcal{D} of 3-tuples (α, λ, P_f) matching. As we know the plausibility of each 3-tuples in \mathcal{D} , we should be able to compute the plausibility of P_t . However, if we use the likelihood principle and set :

$$pl_d^{\mathcal{P}}(P_t) = pl^{\mathcal{T}}(\mathcal{D}) \quad (24)$$

we obtain an inextricable equation. Indeed, after combination, the plausibility function described in equation (23) is not consonant anymore. Then, we cannot derive the plausibility function from the contour function. Hence, we decide to define a consonant plausibility function such that its contour function fulfil this condition:

$$pl_d^{\mathcal{P}}(P_t) = \max_{(\alpha, \lambda, P_f) \in \mathcal{D}} pl^{\mathcal{T}}(\alpha, \lambda, P_f) \quad (25)$$

In this case, this belief function is more specific than the one we would obtain using equation (24). However, as the plausibility of each 3-tuple is built using a singleton P_t of \mathcal{P} , it can be considered like a good trade-off.

The algorithm 2 sums up all the results obtained in section 5 and helps us to compute the plausibility to detect an object at a given distance of the sensor.

Data: cf. table 1

Result: Plausibility function of P_t for a given distance d

initialization;

for $C_i = (\alpha, \lambda, Pf) \in [0 : h : 1] \times [0 : h : 1] \times [0 : h : 1]$ **do**

for $d_j \in \{5, 10, 20\}$ **do**

$P_{t(i,d_j)} = \text{function}(C_i, d_j)$; // cf. equation (14)

$pl_{d_j}^P(P_{t(i,d_j)}) = \text{function}(n_{d_j}, k_{d_j}, P_{t(i,d_j)})$; // cf. algorithm 1

$pl_{d_j}^T(C_i) = pl_{d_j}^P(P_{t(i,d_j)})$;

end

$pl^T(C_i) = \prod_j pl_{d_j}^T(C_i)$;

$P_{t(i,d)} = \text{function}(C_i, d)$;

end

$\mathcal{D} = \{C_i | \text{abs}(P_{t(i,d)} - P_t) \leq h\}$;

$pl_d^P(P_t) = \max_{C_i \in \mathcal{D}} pl^T(C_i)$

Algorithm 2: How to compute plausibility.

6 Results

The results of this work appear in figure 4. The curve $d = 5m$ (respectively $d = 10m$ and $d = 20m$), represents the plausibility that the probability of true positive of the sensor is equal to P_f at the distance $d = 5m$ (respectively $d = 10m$ and $d = 20m$). To build this curve, we have only used data given by the table 1 and the reasoning described in the section 4. To plot the other curves, we have used the sensor model described in the section 3 and the combination and inference process suggest in section 5.

We remark three things. Firstly, the plausibility function of P_t when $d = 10000$ could be linked to the one of P_f as P_t tends to P_f when d increases. Indeed, we have assumed that we use a passive sensor. Secondly, the closer the sensor is of our points of measure, the smaller is the peak of the plausibility function of P_t . Hence in this case the estimate of the probability of detection is more accurate. Thirdly, the plausibility functions look like step-wise functions. It is normal as we found the construction of a plausibility function on the binomial law ($\mathcal{B}(n, k)$) which is discrete.

7 Conclusion

This study has proposed a way to deal with small amount of data in order to estimate the performance of a sensor. Continuous belief functions have been used in two contexts, the first has been to represent an uncertain knowledge about the distribution of probability describing an phenomenon. The second has been to infer information within a parametric model. In this kind of situation, we

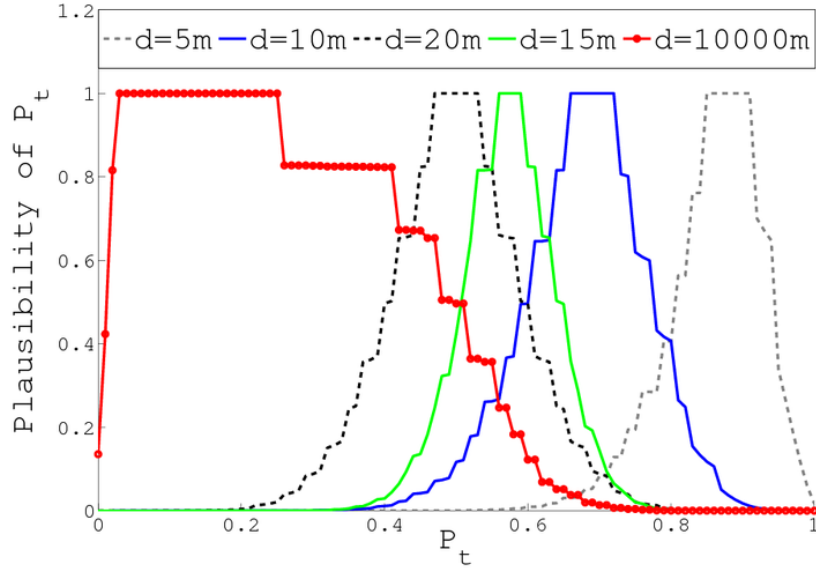


Fig. 4. Plausibilities of P_t for different distances.

are often faced inextricable situation and we have to make our approach more simple to find a solution.

This work is a step to define the coverage a sensor networks. Now we have to study the question a the fusion of the information coming from the different sensors in order to have a better estimation of the global network performance. We have also to include some consideration about the reliability of a sensor to improve the study.

References

1. A. Aregui and T. Denœux, "Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities," *International Journal of Approximate Reasoning*, vol. 49, no. 3, pp. 575–594, 2008.
2. F. Caron, B. Ristic, E. Duflos, and P. Vanheeghe, "Least committed basic belief density induced by a multivariate Gaussian: formulation with applications," *International Journal of Approximate Reasoning*, vol. 48, no. 2, pp. 419–436, 2008.
3. A. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325–339, 1967.
4. A. Dempster, "Upper and lower probabilities generated by a random closed interval," *The Annals of Mathematical Statistics*, vol. 39, no. 3, pp. 957–966, 1968.
5. A. Dempster, "A Generalization of Bayesian Inference," *Journal of the Royal Statistical Society*, vol. 2, no. 30, pp. 205–247, 1968.
6. P.-E. Doré and A. Martin. About using beliefs induced by probabilities. *Workshop on the theory of belief functions (Brest France)*, 2010.

7. D. Dubois and H. Prade, "The principle of minimum specificity as a basis for evidential reasoning," in B. Bouchon & R. R. Yager (Eds.), *Uncertainty in knowledge-based systems*, pp. 75–84, Springer Verlag, Berlin, 1987.
8. D. Dubois, H. Prade, and S. Sandri, "On possibility/probability transformations," in *Proceedings of 4th IFSA Conference*. Citeseer, 1993, pp. 103–112.
9. A. Elfes, "Occupancy Grids: A Stochastic Spatial Representation for Active Robot Perception," in *Autonomous Mobile Robots: Perception, Mapping, and Navigation*, Vol. 1, pp. 60–70, IEEE Computer Society Press, Los Alamitos, CA, 1991.
10. G. Mauris, "Transformation of bimodal probability distributions into possibility distributions," *IEEE Transactions on Instrumentation and measurement*, vol. 59, no. 1, pp. 39–47, 2010.
11. S. Meguerdichian, F. Koushanfar, G. Qu, and M. Potkonjak, "Exposure in wireless ad-hoc sensor networks," in *Proceedings of the 7th annual international conference on Mobile computing and networking*, pp. 139–150, 2001.
12. B. Ristic and P. Smets, "Belief function theory on the continuous space with an application to model based classification," in *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU*, pp. 4–9, 2004.
13. G. Shafer, *A mathematical theory of evidence*. Princeton University Press Princeton, NJ, 1976.
14. P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of Approximate Reasoning*, vol. 9, pp. 1–1, 1993.
15. P. Smets, "Belief functions on real numbers," *International Journal of Approximate Reasoning*, vol. 40, no. 3, pp. 181–223, 2005.
16. P. Smets and B. Ristic, "Kalman filter and joint tracking and classification based on belief functions in the TBM framework," *Information Fusion*, vol. 8, no. 1, pp. 16–27, 2007.
17. P. Smets and B. Ristic, "The combination of evidence in the transferable belief model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 447–458, 1990.
18. T. Strat, "Continuous belief functions for evidential reasoning," in *Proceedings of the 4th National Conference on Artificial Intelligence, University of Texas at Austin*, 1984.